

Er norsk et naturlig språk?

Victoria Rosén

1. Innledning

Det norske språket skiller seg ut internasjonalt sett. Mange land har flere offisielle språk, men Norge er sannsynligvis det eneste land som har to offisielle skriftspråk for det samme språket. Likevel er det kanskje de mange alternative skrivemåtene innenfor både bokmål og nynorsk som skiller norsk mest ut fra andre språk.

Mange som deltar i normeringsdebatten om norsk skriftspråk, mener at den store tillatte variasjonen i offisiell rettskrivning er en verdifull ressurs. Det er mulig at den alminnelige språkbrukeren ser det slik, men jeg kjenner ikke til undersøkelser som viser dette. Denne eventuelle medaljen har i tilfellet også en bakside. Det blir stadig vanligere at naturlige språk brukes i språkteknologiske prosjekter, og norsk har også kommet med i denne utviklingen i de siste årene. For mange typer språkteknologiske produkter er en god ordliste en viktig forutsetning. Slike ordlister er uunnværlige i for eksempel programmer for automatisk korrekturlesning og maskinoversettelse. Men med en gang det er en datamaskin som skal bruke ordlisten, og ikke et menneske, blir den store valgfriheten et problem. Hvis det er mange alternative skrivemåter for et ord, hvordan skal et retteprogram velge et passende forslag for korreksjon? Språkrådet er interessert i at “norsk språk, bokmål og nynorsk, skal sikrast som fullverdig uttrykksmiddel i språkteknologi og annen informasjons- og kommunikasjonsteknologi (IKT)” (Faarlund og Gundersen 1999: 11). Men for at dette skal være mulig, må bokmål og nynorsk ha ordlister som det er hensiktsmessig å bruke i slike sammenhenger. Et menneske som kan norsk, skjønner med en gang at substantivet *mage* er en nøytral form som vil passe i en hvilken som helst tekst, og at *mave* har en markert stilverdi. Men en datamaskin kan ikke skille mellom dem med hensyn til stilverdi.

I denne artikkelen vil jeg først si litt om hvilke typer variasjon som finnes innenfor offisiell rettskrivning for bokmål. I del 3 vil jeg presentere kort noen av hovedtrekkene i de nye vedtakene som Språkrådet har kommet med, som tar sikte på å gjøre en innsats på nettopp dette området. Så vil jeg presentere resultatene av en undersøkelse jeg har gjennomført for å finne ut noe om bruksfrekvensen for en del varianter i moderne bokmål.

2. Variasjonen i bokmål

Det er velkjent at den offisielle rettskrivningen for bokmål rommer alternative skrivemåter for mange ord. Variasjonen gjelder både stammer og bøyningsformer. Noen eksempler på likestilte stammer og bøyningsformer er gitt i henholdsvis (1) og (2).

- | | | | |
|-----|------------------|-----------------|---------|
| (1) | ben | bein | |
| | blek | bleik | |
| | dugg | dogg | |
| | gammel | gammal | |
| | gjemme | gjømme | |
| | hvit | kvit | |
| | høy | høg | |
| | kirke | kjerke | |
| | mave | mage | |
| | sogn | sokn | |
| | spille | spelle | |
| | utenriksminister | utariksminister | |
| (2) | bød | bydde | baud |
| | dronningen | dronninga | |
| | enset | enste | ensa |
| | fløy | flaug | |
| | forledet | forleda | |
| | formanet | formante | formana |
| | gjess | gjæser | |
| | kuer | kyr | |
| | laget | lagd | laga |
| | lavet | lavde | lava |

områdene	områda	
opphøyede	opphøyete	opphøya

I tillegg til at man kan ha likestilte former, finnes det et system med sideformer (også kalt klammeformer). Sideformene er en del av den godkjente rettskrivningen, men de kan ikke brukes i læreboknormalen. Sideformer finnes både for stammer og for bøyning. Eksempler på sideformer og hovedformer av henholdsvis stammer og bøyninger gis i (3) og (4).

(3)	[antagelig]	antakelig		
	[bonn]	botn	bunn	
	[deltagelse]	deltakelse	deltaking	deltaing
	[flyve]	fly	flyge	
	[kor]	hvor		
	[korfor]	[åffer]	hvorfor	
	[leier]	leder		
	[raud]	rød		
	[røke]	røyke		
	[sløkke]	slokke	slukke	
	[viss]	hvis		
(4)	[barnene]	barna		
	[bedro]	bedrog		
	[bestreid]	bestrei	bestridde	bestred
	[greidde]	greide		
	[komma]	komme		
	[kuen]	kua		
	[tigerer]	tigere	tigrer	

Når det finnes så mange variasjonsmuligheter både i henhold til valg av stamme og til valg av bøyningsmåte, sier det seg selv at det blir svært mange kombinasjoner når former med varianter inngår i sammensetninger. Her er bare ett eksempel fra *Bokmålsordboka*, *handslokkingsapparat*, som har følgende ti alternative skrivemåter:

- (5) handslokkingsapparat
handslokningsapparat
handslukningsapparat
handslukningsapparat

handslökkingsapparat
 håndslokkingsapparat
 håndslokningsapparat
 håndslukkingsapparat
 håndslukningsapparat
 håndslökkingsapparat

Formene der stammen *sløkk*- inngår, er sideformer, alle andre er likestilte former. Siste sammensetningsledd, *apparat*, har alternative bøyningsformer i ubestemt flertall (*apparat* eller *apparater*) og i bestemt flertall (*apparata* eller *apparatenene*). Selv om man koder formene som ender på *apparat* som enten entall eller flertall, vil man ende opp med et paradigme med 50 ordformer for dette ene ordet.

En del av oppslagene i *Bokmålsordboka* står i parentes. Det betyr at de har såkalt ikke-godkjent skrivemåte, og slike oppslag har henvisning til en form med godkjent skrivemåte. Eksempler på henholdsvis ikke-godkjente og godkjente former vises i (6).

(6)	(grube)	gruve	
	(guerilla)	gerilja	
	(have)	hage	
	(hverken)	verken	
	(hvirvel)	virvel	
	(høl)	hull	
	(høytflyvende)	høgtflygende	høytflygende
	(jasmin)	sjasmin	
	(sne)	snø	
	(syv)	sju	
	(tyve)	tjue	

Vekslingen mellom former som *hverken* og *verken* kan altså ikke sies å være en del av variasjonen i den offisielle rettskrivningen, selv om den i høy grad, som vi senere vil se, er en del av den variasjonen som finnes i bokmålstekster.

Det finnes enda en type variasjon som tilhører offisiell rettskrivning, men denne typen er ikke så godt synlig i *Bokmålsordboka* som variasjonen mellom likestilte hovedformer og sideformer. Det er nemlig en god del oppslag der et ord står oppført alene, men med et annet ord (eller i noen tilfeller, flere andre ord) som sin definisjon. Dette gjelder for eksem-

pel ordene i venstre kolonne i eksempel 7; ordene i høyre kolonne er blant de ord som står oppført som definisjoner for disse.

(7)	annleis	annerledes
	ange	lukt
	dåm	duft
	handsame	behandle
	hugnadsam	hyggelig
	høve	anledning
	høve	passe
	kjæte	kåthet
	vonbroten	skuffet
	åkkesom	uansett

I noen tilfeller ligner ordene på hverandre formmessig, som for eksempel *annleis* og *annerledes*, mens det i andre tilfeller er ord som klart ikke er etymologisk relatert til hverandre.

Noen av formene som den offisielle rettskrivningen rommer, virker så fremmede for bokmål at man kan lure på om noen overhodet bruker dem. Helge Omdal har gjennomført en undersøkelse blant grunnfagsstudenter i nordisk der siktemålet var å finne ut noe om det han kaller språknormsikkerhet, altså i hvilken grad personer med utdanning utover obligatorisk skolegang kjenner til hva slags former som hører til rettskrivningen. Forsøkspersonene fikk presentert konstruerte tekster på bokmål og nynorsk, og de skulle "rette" disse tekstene i samsvar med det de oppfattet som korrekt språk. Bokmålsteksten gjengis nedenfor (Omdal 1996: 166).

- (8) Ola sitter på en sten og venter på de to fiskerene.
 Katja ser på han og sier; «Må du alltid vente på de? Kan ikke de komme til deg nå i haust?»
 At Ola ikke likte disse spørsmåla er tydelig.
 Han ser på ho gjennom brilleglassa og trekker frem et papir.
 De er framdeles på talefot, når det gjelder allmene ting.

Her er en del feil som bør korrigeres. Men det er også en del uvanlige former som hører inn under rettskrivningen, og som dermed ikke bør korrigeres; disse formene er tatt med nettopp fordi Omdal er usikker på om studentene vet at de er tillatt.

Omdal kommenterer bare studentenes svar på tre av ordene i teksten: *fiskerene*, *haust* og *spørsmåla*. Så mange som en tredjedel av studentene lot være å rette formen *fiskerene*. Omdal mener at en del av dem som ikke rettet formen, kan ha kjent til riktig skrivemåte, men at de ganske enkelt har vært uoppmerksomme og har oversett feilen. Likevel mener han at det må være slik at de fleste som ikke har foretatt denne rettelsen, ikke har visst at *fiskerne* er den eneste formen i bestemt form flertall for dette ordet. Han mener også at usikkerhet om hvorvidt det skal være sammentrekning, kan henge sammen med manglende konsekvens i normalen, og peker på vekslingen mellom skrivemåter som *filteret* og *filtret*, *fortvilet* og *fortvilt*.

De to andre kommenterte ordene, *haust* og *spørsmåla*, er korrekte i henhold til rettskrivningen. Det er interessant at 94 % av studentene rettet *haust* til *høst*; de var altså ikke klar over at *haust* er sideform, og dermed tillatt å skrive i bokmål. Enda mer interessant er det at så mange som 65 % rettet *spørsmåla* til *spørsmålene*. Omdal kommenterer dette resultatet på følgende måte (Omdal 1996: 169–170):

Så godt som alle intetkjønnsord i bokmål kan ha endinga *-a* ved sida av *-ene* i bestemt form flertall. Dette er med andre ord et sentralt og gjennomgående trekk ved bokmålets bøyningssystem, og resultatet [...] er derfor mildt sagt forbausende. Dersom en også her rekner med enkelte uoppmerksomme studenter, har mer enn to tredeler meint at *a*-endinga i *spørsmåla* ikke er i samsvar med gjeldende rettskriving for bokmål, og derfor korrigert forma til *spørsmålene*. Hva grunnen til denne mangelen på det vi kan kalle elementær grammatisk kunnskap, kan være, skal vi ikke spekulere over her. Vi skal bare antyde at det trulig kan ha sammenheng både med opplæring og bruksfrekvens i skriftspråklig sammenheng, f.eks. står *-a*-formene i bokmål fortsatt svakt i de fleste avisene, og i ukepresse, bøker (også lærebøker) og film- og videotekster. Likevel er det et merkbart sprang fra å oppfatte visse bøyningsformer som mindre utbredt enn andre, og til å oppfatte dem som feil.

Det er riktig at vekslingen mellom *-a* og *-ene* er “et gjennomgående trekk ved bokmålets bøyningssystem”, hvis man med bøyningssystem mener de vedtak som er gjort med hensyn til hva som er tillatt å skrive innenfor offisiell rettskrivning. Men det er ikke dermed sagt at det er et sentralt og gjennomgående trekk ved bokmål slik det faktisk skrives av dem som bruker det. Omdal har sikkert rett når han antyder muligheten for at lav bruksfrekvens i skriftlig bokmål kan være en del av forklaringen på at studenter ikke vet at det er ‘lov’ å skrive *spørsmåla*. Hvor lite frekvent denne bøyningsmåten er, kommer vi tilbake til i del 4.

3. Språkrådets forslag til vedtak

Som sagt i innledningen er det potensielt et stort problem for automatisk språkbehandling at det finnes så mange alternative former i de offisielle ordlistene. En opplagt løsning på problemet er å fjerne former som ikke er i bruk. Språkrådet har satt i gang et slikt arbeid. I flere år har fagnemnden for bokmål arbeidet med “å ta bort former som har vært tillatt i lengre tid, men som i liten grad er blitt tatt i bruk av dem som skriver bokmål” (Kulbrandstad 1999: 1). De fleste av de endringene nemnden har foreslått, ble vedtatt av Språkrådet i februar 2000, og vedtakene ligger, når dette skrives, til godkjenning i Kulturdepartementet.

Den viktigste endringen er at nemnden foreslår at systemet med sideformer skal legges ned.¹ I en pressemelding fra Språkrådet (<http://www.sprakrad.no/2000bm.htm>) gis det følgende eksempler på hva slags sideformer som dermed går ut av rettskrivningen (min kursivering):

De som forsvinner, vil være former og ord som er lite brukt i skriftlig bokmål. Det gjelder for eksempel former som [*feierer*], [*lærer*] og [*taler*], adjektivformer med dobbeltkonsonant etter diftong av typen [*breitt*] og [*flautt*] og ca. 60 enkeltformer som i liten grad har blitt tatt i bruk i bokmål, blant andre [*berr*], [*fekk*], [*frå*], [*ho*], [*hvems*], [*korfor*], [*måndag*], [*åffer*] og [*såg*].

Men det er ikke bare slik at sideformer forsvinner; en god del former som til nå har vært sideformer, blir gjort om til hovedformer. Dette gjelder bl.a. noen få intetkjønnsord som tidligere har hatt *-ene* som sideform i bestemt form flertall. Tidligere kunne man bare skrive f.eks. *barna* og *beina* i læreboknormalen, men nå blir *barnene* og *beinene* likestilt med disse. En lignende forandring gjelder drøyt 900 substantiv som siden 1981 har hatt *-a* som hovedform i bestemt form entall og *-en* som sideform. Altså, former som *bruen*, *jenten* og *kjerringen* blir nå likestilt med *brua*, *jenta* og *kjerringa*. Disse endringene innebærer at det nå er full valgfrihet mellom *-a* og *-en* i bestemt form entall av hunkjønnsst substantiv, og mellom *-a* og *-ene* i bestemt form flertall av intetkjønnsst substantiv.

Det er også foreslått mange mer detaljerte endringer. Men det mest interessante er den gjennomgripende endringen med hensyn til sideformsystemet. I det følgende vil jeg ta opp forhold som er relatert til denne endringen.

4. En undersøkelse om bruken av valgfrie former i bokmål

Jeg har brukt Oslo-korpuset av taggedede norske tekster til å undersøke forekomsten av noen av de alternative formene som finnes i bokmål. Bokmålsdelen av korpuset (<http://www.tekstlab.uio.no/norsk/bokmaal/>) inneholder ca. 18,5 millioner ord fordelt på avistekster, sakprosattekster og skjønnlitterære tekster. Korpuset innebærer en helt ny mulighet for å undersøke aspekter ved moderne norsk skriftspråk på en rask og effektiv måte.²

Mange av de forandringene som Språkrådet har foreslått, gjelder tidligere sideformer som er gjort om til hovedformer. En av disse forandringene er at ingen intetkjønnssubstantiv lenger skal ha *-a* som hovedform og *-ene* som sideform; disse endelsene skal nå være likestilte. Jeg har undersøkt forekomsten av begge formene for de ni ordene som tidligere hadde obligatorisk *-a*, og resultatene står i tabell 1.

avhola	0	avholene	0
barna	2212	barnene	4
beina	392	beinene	0
beista	0	beistene	3
hola	0	holene	0
kola	0	kolene	0
krøttera	8	krøtterne	0
nauta	0	nautene	1
segla	1	seglene	1

Tabell 1 Fordeling av *-a* og *-ene* i substantiv som tidligere hadde obligatorisk *-a*

Denne forandringen ser ikke ut til å få de store konsekvenser, siden så godt som alle bokmålsbrukere allerede bruker *a*-endelsen i de eneste to av disse ordene som er noe særlig i bruk. Men Språkrådet har heller ikke ment at det ville få store konsekvenser; poenget med vedtaket var heller å slippe å ha unntak fra en bøyningsmåte som ellers gjelder nesten alle intetkjønnssubstantiv.

Som nevnt ovenfor påpekte Omdal at *a*-endelsen i bestemt form flertall av intetkjønnssubstantiv står svakt. For å undersøke hvor vanlig denne bøyningsmåten er, har jeg søkt på 40 substantiv. Den relative frekvensen av endelsene *-a* og *-ene* for disse ordene fremgår av tabell 2.

arbeida	0	arbeidene	140
beita	0	beitene	38
beløpa	0	beløpene	65
ben	29	benene	43
bestanda	0	bestandene	25
bilda	0	bildene	351
brudda	0	bruddene	24
bånda	0	båndene	39
dyra	105	dyrene	305
emna	0	emnene	41
epla	0	eplene	38
erma	0	ermene	26
faga	2	fagene	245
flya	0	flyene	79
glassa	0	glassene	58
hoda	0	hodene	102
husa	14	husene	240
kontora	0	kontorene	67
ledda	0	leddene	38
lokala	0	lokalene	223
låna	0	lånene	121
menneska	0	menneskene	810
merka	0	merkene	42
møta	0	møtene	187
områda	0	områdene	944
organa	0	organene	64
regima	0	regimene	22
rykta	0	ryktene	70
seila	0	seilene	24
smykka	0	smykkene	19
spørsmåla	1	spørsmålene	439
stykk	0	stykkene	33
styra	0	styrene	24
teppa	0	teppene	17
tilbuda	0	tilbudene	153
tilfella	0	tilfellene	361
vitna	0	vitnene	50
værelsa	0	værelsene	10
yrka	0	yrkene	16
åra	85	årene	4307

Tabell 2 Fordeling av *-a* og *-ene* i en del intetkjønnsstavn

Jeg ble svært overrasket over hvor få ord som overhodet hadde noen forekomster av *a*-endelsen. Ordene i tabellen er tilfeldig valgt; jeg bladde gjennom listen av intetkjønnsstavn³ og søkte på ord der jeg mente det var en sjanse for å finne forekomster (noen ord er så lite frekvente at de ikke forekommer med noen av endelsene, jf. tabell 1). Men det er i hvert fall ikke slik at jeg valgte ord for å favorisere *-ene*. Tvert imot prøvde jeg så godt jeg kunne å finne *a*-endelser, og jeg måtte faktisk holde på ganske lenge før jeg fant noen i det hele tatt. Det ser ut til at det er bare noen få ord der (noen) bokmålsbrukere finner *a*-endelsen naturlig.

Den andre forandringen som Språkrådet foreslår når det gjelder *a*-endelser, er at *-en* ikke lenger skal være sideform for stavn som tidligere har hatt obligatorisk *-a* i bestemt form entall. I pressemeldingen fra Språkrådet står følgende (min kursivering):

Siden 1981 har drøyt 900 hunkjønnsord hatt *-a* som hovedform i bestemt form entall og *[-en]* som sideform, for eksempel *brua*, *[bruen]*. Etter forslaget vil alle disse ordene få jamstilte former med *-a* og *-en*. Det gjelder frekvente ord som *bru*, *brygge*, *bukt*, *dyng*, *flis*, *grøft*, *hytte*, *jente*, *kjerring*, *kone*, *ku*, *myr* og *øy*. Selv om en etter forslaget altså vil kunne skrive for

eksempel *jenten* og *kjerringen*, er det ikke Språkrådets intensjon med endringen å støtte bruken av slike former. De fleste av disse formene forekommer sjelden eller aldri i bokmålstekster, og det forventes heller ikke at de vil bli vesentlig oftere brukt dersom den foreslåtte endringen blir vedtatt. Begrunnelsen for forslaget er ønsket om en systemforenkling ved at en får en enkel regel uten unntak, og at det blir enklere å bruke ordlistene.

For å undersøke påstanden om at de fleste av formene på *-en* sjelden eller aldri forekommer i bokmålstekster, har jeg søkt på disse ordene (og på *bro*, som er likestilt med *bru*) i bestemt form entall, både med *a*-endelse og med *en*-endelse. I tillegg har jeg søkt på dem i ubestemt form umiddelbart etter ubestemt artikkel i maskulinum og femininum. Resultatene står i tabell 3.

broa	0	broen	262	ei bro	0	en bro	290
brua	38	bruen	0	ei bru	4	en bru	0
brygga	17	bryggen	40	ei brygge	0	en brygge	5
bukta	12	bukten	20	ei bukt	0	en bukt	3
dynga	3	dyngen	0	ei dyng	0	en dyng	0
flisa	5	flisen	6	ei flis	0	en flis	5
grøfta	16	grøften	33	ei grøft	5	en grøft	16
hytta	151	hytten	142	ei hytte	8	en hytte	41
jenta	290	jenten	125	ei jente	71	en jente	163
kjerringa	49	kjerringen	5	ei kjerring	6	en kjerring	3
kona	280	konen	146	ei kone	3	en kone	29
kua	73	kuen	5	ei ku	13	en ku	11
myra	275	myren	13	ei myr	1	en myr	11
øya	249	øyen	71	ei øy	1	en øy	22

Tabell 3 Fordeling av maskulinum og femininum for substantiv som tidligere hadde obligatorisk *-a* i bestemt form entall

Som tabellen viser, er det ikke tilfellet at formene på *-en* forekommer sjelden eller aldri. Når det gjelder *brygge*, *bukt*, *flis* og *grøft*, er det faktisk slik at formen med *-en* forekommer oftere enn den med *-a*. Og selv om det er flere forekomster av *hytta*, *jenta* og *kona* enn av henholdsvis *hytten*, *jenten* og *konen*, er formene med *-en* sterkt representert. Bare to av formene, *bruen* og *dyngen*, har ingen forekomster i korpuset. Når det gjelder *dyngen*, kan dette skyldes at ordet i det hele tatt er så lite frekvent. Forklaringen må være en annen for *bruen*. Her ser vi en helt klar fordeling: *bro* blir behandlet som hankjønnssubstantiv og *bru* som hunkjønnssubstantiv, både i bestemt og ubestemt form.

Det er interessant at hunkjønnsformene står så mye svakere for disse ordene enn pressemeldingen fra Språkrådet synes å anta. Det ville være interessant å vite hva en påstand som “De fleste av disse formene forekommer sjelden eller aldri i bokmålstekster” er basert på. Det bør være velkjent at trekjønnsystemet står svakt i bokmål. Dette ble påpekt av Einar Haugen så tidlig som i 1966. Noen få ord forekommer hyppigere med *a*-endelse enn med *-en*, men generelt sett står nok de tradisjonelle riksmålsformene sterkest. Og når det gjelder kongruens med andre ord, står femininum enda svakere. Hvis vi ser bort fra *bru*, som ser ut til å være en ren hunkjønnsform, er det bare to av ordene i listen ovenfor som forekommer oftere med *ei* enn med *en*, nemlig *kjerring* og *ku*. Så lenge bruken av hunkjønn stort sett er begrenset til en bøyningsform av substantiv, kan man ikke snakke om et ekte trekjønnsystem. Peter Trudgill har kalt systemet i bokmål et system med 2,5 kjønn, og det er ganske treffende for mange bokmålsbrukere, som bruker en del *a*-endelser, men ikke andre ord som kongruerer i hunkjønn, som artikler og adjektiv.

Det finnes bare to adjektiv, *eigen* og *liten*, som har egne hunkjønnsformer i bokmål. I tillegg har *annen* de likestilte formene *annen* og *anna* i hunkjønn. Søk på ubestemt artikkel fulgt av disse tre ord gir resultatene vist i tabell 4.

ei eiga	0	en eigen	0
ei lita	14	en liten	1709
ei anna	9	en annen	2826

Tabell 4 Fordeling av hankjønns- og hunkjønnsformer av *annen*, *eigen* og *liten*

Som vi ser, er de feminine formene svært lite brukt. Søk på ordsekvensen *en egen* viser at den forekommer 799 ganger. Formen *eigen* finnes faktisk overhodet ikke i korpuset, og formen *eiga* forekommer bare en gang, i frasen *min eiga studietid*, som neppe kan regnes som korrekt bokmål.

Hittil har vi sett på forhold relatert til forslagene til forandringer i substantivbøyningen. Språkrådet foreslår også en del forandringer når det gjelder bøyning av verb. Et eksempel er at preteritumsformer på *[-eid]* i verb av typen *gli* blir gjort om til hovedformer. Jeg har undersøkt forekomsten av alle preteritumsformer for verbene dette gjelder, og resultatet vises i tabell 5.

bestred	4	bestrei	0	bestreid	0	bestridde	0
gled	237	glei	8	gleid	0		
gned	47	gnei	1	gneid	0	gnidde	4
led	12	lei	1	leid	0	lidde	0
red	53	rei	5	reid	0		
skled	34	sklei	3	skleid	0	sklidde	1
stred	9	strei	2	streid	0	stridde	3
sved	37	svei	7	sveid	0	svidde	2
vred	70	vrei	6	vreid	0	vridde	5

Tabell 5 Fordeling av preteritumsformer for verb av typen *gli*

Som tabellen viser, står den sterke formen med monoftong sterkest for alle disse verbene. Det er også noen få forekomster av former med diftong uten *d* og av svak form, men ikke en eneste forekomst av formene som slutter på *-eid*. Det er derfor vanskelig å forstå hvorfor man har foreslått at disse formene skal gjøres om fra sideformer til likestilte hovedformer. Dette resulterer i så mange som fire likestilte former for noen av verbene. Siden formene på *-eid* ser ut til å stå så svakt, ville det ha vært mer i tråd med intensjonen om å fjerne former som er lite brukt, om de hadde vært tatt ut av rettskrivningen.

Verbet *bli* har hatt følgende bøyning i preteritum: *ble* og *blei* har vært likestilte hovedformer, mens *vart* har vært sideform. Språkrådet foreslår at *vart* skal gå ut. Søk på disse formene gir resultatene vist i tabell 6.

ble	45141	99,3 %
blei	301	0,7 %
vart	0	0,0 %

Tabell 6 Fordelingen av preteritumsformer av *bli*

Det er ingen forekomster av *vart* i korpuset, så forslaget om at denne formen skal tas ut av rettskrivningen, virker fornuftig. Samtidig innebærer forslaget at *blei* skal forbli likestilt med *ble*. Som vi ser av tabellen, står *blei* for bare 7 promille av forekomstene. Her kan det nevnes at 256 av de 301 forekomstene av *blei* finnes hos to forfattere, 70 ganger hos Sverre Fredheim (i romanen *Som to blåklommer, sa hun*) og 186 ganger hos Hans Prestbakmo (i *NOU 21 1994*).

Et annet eksempel der en sideform er foreslått fjernet mens to likestilte hovedformer står igjen, gjelder substantivet *bunn*, som har hatt den likestilte formen *botn* og sideformen *bonn*. Fordelingen av disse formene er gitt i tabell 7.

bunn	261	97,0 %
botn	5	1,9 %
bonn	3	1,1 %

Tabell 7 Fordelingen av *bunn*, *botn* og *bonn*

Selv om *bonn* har vært sideform og *botn* har vært hovedform, er de omtrent like lite brukt.

Eksempelene jeg har vist så langt, viser klart at problemet med lite brukte former i rettskrivningen ikke begrenser seg til sideformer. Både når det gjelder bøyning, slik som med *spørsmålene* og *spørsmåla*, og når det gjelder stammer, som *bunn* og *botn*, kan den ene formen være svært lite brukt på tross av at den er en likestilt hovedform. Dette ser vi også med ord som i eksempel (7) ovenfor. Forekomstene av noen av disse ordene er vist i tabell 8.

vettug	4	fornuftig	383
hugnadsam	0	hyggelig	562
vonbroten	1	skuffet	552
handsame	0	behandle	640

Tabell 8 Fordeling av radikale og tradisjonelle synonyme former

I sterk kontrast til dette er bruken av en del av de ikke-godkjente ordene vist i eksempel (6).

verken	984	hverken	755
snø	423	sne	83
sju	1067	syv	750
tjue	161	tyve	204

Tabell 9 Fordeling av godkjente og ikke-godkjente former

Her er det interessant å merke seg at *sne* er betydelig sjeldnere enn *sno*, mens de andre formene er omtrent like vanlige. Dette stemmer godt overens med deres stilverdi; mens *sne* har en klart konservativ valør, så virker *hverken*, *syv* og *tyve* ganske nøytrale. De står sterkt i skriftlig bokmål, til tross for at de har vært 'forbudt' i flere tiår.

Som vist i del 2 ovenfor, kan det bli veldig mange muligheter når ledene i en sammensetning har varianter. Jeg viser resultat av søk for noen ganske frekvente sammensetninger av denne typen i tabell 10.

helstøpt	21	helstøpt	0	heilstøpt	0	heilstøpt	0
hvetemel	119	hvetemjøl	0	kveitemel	0	kveitemjøl	0
hvitløk	130	hvitlauk	0	kvitløk	0	kvitlauk	0
meningsløs	61	meningslaus	0	meiningsløs	0	meiningslaus	0
sykehjem	503	sykeheim	0	sjukehjem	0	sjukeheim	26

Tabell 10 Fordeling av alternative sammensetningsformer

Her må det bemerkes at 25 av de 26 forekomstene av *sjukeheim* var som del av navn, for eksempel *Rennesøy sjukeheim*.

Når man har alternative stammer for et ord, hender det ofte at begge utgjør første eller andre ledd i en lang rekke sammensetninger. Vi skal se på paret *brudd/brott*. I tillegg til å stå som enkeltord, står begge formene som alternative siste ledd i 48 sammensetninger. I tillegg til sammensetningene med *-brudd* har vi altså følgende:

- (9) armbrutt, avbrutt, avtalebrutt, beinbrutt, brevbrutt, dagbrutt, diskresjonsbrutt, ekteskapsbrutt, forfatningsbrutt, frambrutt, følesesutbrutt, gjennombrutt, gledesutbrutt, hellebrutt, innbrutt, kalkbrutt, kontraktbrutt, kraniebrutt, leddbrutt, lårbeinsbrutt, løftebrutt, linjebrutt, lovbrutt, nervesammenbrutt, oppbrutt, ribb-beinsbrutt, ribbeinsbrutt, ribbensbrutt, saltbrutt, sammenbrutt, skiferbrutt, skipbrutt, skybrutt, steinbrutt, stemningsutbrutt, stilbrutt, strømbrott, svangerskapsavbrutt, tillitsbrutt, traktatbrutt, tretthetsbrutt, trøtthetsbrutt, troskapsbrutt, truskapsbrutt, utbrutt, utmatningsbrutt, vennskapsbrutt, vulkanutbrutt

Det ville være svært tidkrevende å søke på hvert sammensetningspar manuelt. Men man kan velge å søke på alle ordene som ender på en viss

bokstavstreng. Denne typen søk finner alle ord som ender på den strengen man oppgir, uansett om noe annet står foran. Altså søk på *-brott* gir både forekomster av *brott* som enkeltord og alle ord som ender på *-brott*. Resultatet av denne typen søk på ord som ender på *-brott*, var ni ord. Det er to forekomster av *brott* som enkeltord:

(10) Alt kom så *brott*.

Han har alltid manøvrert sikkert i *brott* og *stille*.

Den første er opplagt en stavfeil, mens den andre er *brott* i betydningen 'brenning', som ikke har noe med vekslingen mellom *brott* og *brudd* å gjøre. Blant de andre forekomstene finner vi *grunnbrott* og *skumbrott* (begge igjen i betydningen 'brenning'), to forekomster av ordet *nybrott*, som ikke har noe tilsvarende stamme på *-brudd*, og endelig tre forekomster av *Kjelbrott* (fotballtreneren Sven). Dette viser at formen *brott* som alternativ til *brudd* ikke forekommer en eneste gang i bokmålsdelen av Oslo-korpuset, hverken som enkeltord eller som del av noen av de 48 sammensetningene med *-brott* som siste ledd som listes opp i *Bokmålsordboka*. Til sammenligning er det totalt 2271 forekomster av formen *brudd*; som enkeltord forekommer det 728 ganger, mens 1543 sammensetninger har *-brudd* som siste del.

5. Konklusjon

Hva er så svaret på spørsmålet som stilles i tittelen på denne artikkelen? Svaret er selvfølgelig 'ja' hvis man forstår 'naturlig språk' i sin vanlige betydning, som fagterm. Men jeg har prøvd å vise at svært mange av de alternative skrivemåtene som finnes for bokmål, ikke er i bruk. Det at et språks offisielle rettskrivning inkluderer mange ord som aldri blir brukt i skriftspråket, kan neppe kalles naturlig. Det kan heller ikke sies å være naturlig at det skal være 'forbudt' å bruke former som *hverken*, som har en godt etablert plass i en ubrutt skriftradisjon.

En innvending som kanskje kan rettes mot denne undersøkelsen, er at korpuset som er brukt, er for lite og/eller for lite representativt til at resultatene kan være pålitelige. Selv om jeg mener at Oslo-korpuset er en god og viktig ressurs for forskning om norsk språk, er det opplagt at vi i Norge trenger større og enda mer representative korpora. Prosjekter som tar sikte på å utvikle slike korpora, er på trappene. En annen mulig innvending er at

selve undersøkelsen er for snever eller for lite representativ, for eksempel at for få ord er tatt med. Jeg har ikke selv hatt anledning til å gjennomføre en større undersøkelse, men jeg mener at det er på tide at en slik undersøkelse blir satt i gang, slik at fremtidig språknormeringsarbeid kan være fundert på empiriske funn. Men jeg tviler sterkt på at en større undersøkelse basert på et større korpus ville komme frem til vesentlig andre resultater.

Det store arbeidet som Språkrådet har gjennomført de siste årene med å fjerne former i bokmål som er lite brukt, er prisverdig. Fjerning av sideformene er et viktig skritt i retning av et mer oversiktlig rettskrivningssystem. Men som jeg har prøvd å vise i denne artikkelen, er det langt fra tilfellet at det bare er blant sideformene man finner lite brukte former. Skal man lykkes med å få slike former ut av ordlistene, må man også se på andre typer variasjon. Det er å håpe at Språkrådet vil fortsette og intensivere dette arbeidet i tiden fremover, slik at man kan komme frem til en offisiell rettskrivning som avspeiler språket slik det skrives av bokmålsbrukerne.

Noter

1. Forslaget fra Språkrådet går rent formelt ut på at systemet skal “stilles i bero” og ikke brukes inntil videre. Dette betyr at man kan slutte å operere med sideformer uten å måtte få forandringen godkjent av Stortinget.
2. Selv om korpuset er et godt verktøy, må det brukes med fornuft. Det er viktig å være klar over at det er en god del nynorsk i bokmålsdelen av korpuset, stort sett på grunn av at noen aviser, som for eksempel *Bergens Tidende*, bruker både bokmål og nynorsk, og det har ikke lyktes å luke ut alle tekster som er skrevet på nynorsk. Dette utgjør en alvorlig feilkilde, så mange ganger må man gå gjennom søkeresultatet manuelt. Det er også en del tekster som hverken er skrevet på bokmål eller nynorsk; her er Alf Prøysens *Trost i taklampa* det klareste eksempelet. Jeg har ikke telt med forekomster når den omgivende konteksten klart ikke kan klassifiseres som bokmål.
3. Listen jeg har brukt, er en av ordlistene som ble utviklet i forbindelse med SCARRIE, et EU-finansiert forskningsprosjekt for automatisk korrekturlesning for dansk, norsk og svensk. SCARRIE-ordlistene er utviklet på grunnlag av materiale fra NorKompLeks, som i sin tur er basert på materiale fra *Bokmålsordboka*.

Referanser

- Faarlund, Jan Terje og Dag Gundersen, red. 1999. *Språknytt* 27: 3–4.
- Haugen, Einar. 1966. *Language Conflict and Language Planning*. Cambridge, Massachusetts: Harvard University Press.
- Kulbrandstad, Lars Anders. 1999. Ikke lenger hovedformer og sideformer i bokmål? *Språknytt* 2: 1–4.
- Landrø, Marit Ingebjørg og Boye Wangensteen, red. 1997. *Bokmålsordboka*. Oslo: Universitetsforlaget.
- Omdal, Helge og Lars Vikør. 1996. *Språknormer i Norge. Normeringsproblematikk i bokmål og nynorsk*. Landslaget for norskundervisning (LNU) og Cappelen Akademisk Forlag as.
- NorKompLeks: <http://benoni.hf.ntnu.no/nkl/>
- Oslo-korpuset av taggede norske tekster: <http://www.tekstlab.uib.no/norsk/bokmaal/>
- SCARRIE: <http://fasting.hf.uib.no/scarrie/>